JOURNAL OF MULTIVARIATE ANALYSIS 39, 217–235 (1991)

Estimating a Survival Function with Incomplete Cause-of-Death Data

SHAW-HWA LO

Columbia University and Harvard University

Communicated by the Editors

We propose a random censorship model which permits uncertainty in the cause of death assessments for a subset of the subjects in a survival experiment. A nonparametric maximum likelihood approach and a "self-consistency" approach are considered. The solution sets corresponding to both approaches are found. They are infinite and identical. Only some of the solutions are consistent; i.e., the MLEs and self-consistent estimators are not consistent in general. Two estimates are thus proposed and their asymptotic properties are studied. It is shown that both estimates are strongly consistent and converge to Gaussian processes. The covariance structures of these Gaussian processes are derived.

1. INTRODUCTION

Let (X_i, Y_i, Δ_i), $i = 1, 2, ..., n$, be i.i.d. random variables, with X_i, Y_i, and Δ_i independent for each i. Let F and G denote the X population and the Y population, respectively. The distribution function (d.f.) F is known as the survival distribution and G is known as the censoring distribution. Let $\Delta_1, \Delta_2, ..., \Delta_n$ denote i.i.d. Bernoulli random variables with common probability $P(\Delta_1 = 1) = p$, and $P(\Delta_1 = 0) = 1 - p$. Finally, let δ_i denote the usual indicator function, i.e.,

$$\delta_i = 1 \quad \text{iff} \quad \{X_i \leqslant Y_i\}$$
$$\delta_i = 0 \quad \text{iff} \quad \{X_i > Y_i\}.$$

Let $Z_i = \min(X_i, Y_i) = X_i \wedge Y_i$.

Suppose that one observes either (Z_i, δ_i) or $(Z_i, \cdot)$ depending on $\Delta_i = 1$ or $\Delta_i = 0$, respectively. Since Y_i and Δ_i are independent of X_i, the random

Received May 1989; revised January 20, 1991.

AMS 1980 subject classifications: 62G05, 62G30.

Key words and phrases: random censorship model with uncertainty, survival experiment, nonparametric maximum likelihood, self-consistency, strongly consistent, Gaussian process.

0047-259X/91 $3.00

variables Y_i and Δ_i may be called noninformative censoring [8] and noninformative uncertainty, respectively.

The second type of data $(Z_i, \cdot)$ simply says that $Z_i = X_i \wedge Y_i$ is observed but δ_i is missing (unknown). This model arises in certain animal experiments, for example, in carcinogenesis bioassay. In this assay a chemical is administered at a constant daily dose rate for the lifetime of the test animals. For each animal in a given experimental dose group, the age-at-death and the presence or absence of specific tumor types are recorded. In the usual statistical approach to bioassay data, each animal must be classified at death as either tumor-free, dead due to the tumor, or dead with the tumor present but not considered responsible for causing death. Although knowing each animal's cause of death simplifies the statistical analysis, most pathologists assert that incidental and fatal occurrences cannot always be distinguished; in fact, forced cause-of-death data can create substantial biases (see [9, 12, 4, 5]).

The model described above permits uncertainty in the cause of death assessments for a subset of the subjects in a survival experiment. In order to avoid problems of nonidentifiability, we assume that the uncertainty occurs randomly (independent of time). An excellent discussion of this issue is given by Lagakos [8]. In the above model we have chosen, the probability of observing δ_i to be a common unknown constant. The arguments given in this paper extend with some additional effort to the case where every subject may have its own distinct p value. However, in order to keep the presentation simple it is assumed here that p is a constant. Although the current work is motivated by data arising from animal experiments, the model and the methods described here may well apply to many studies of disease in patients when the cause of death is not certain.

Quality control provides another area of possible application. This, in particular, if it is not always certain that the failure of a system is due to the failure of a specific component of the system.

Nonparametric MLEs have been widely used in survival analysis. Another promising method of estimating survival functions with incomplete data is the "self-consistency" approach introduced by Efron [6]. In Section 2 we shall consider both approaches. Our findings indicate the following: (i) Neither approach leads to a unique solution. In fact, the solution sets corresponding to these two approaches are infinite. Furthermore, it is shown that these solution sets are identical. (ii) Only some of the solutions would lead to consistent estimators, i.e., the MLEs and self-consistent estimators are not consistent, in general. This unpleasant phenomenon is a consequence of the fact that both approaches might ignore information provided by some of the data that are present. In Section 3 we propose two new estimators. Their asymptotic properties are studied. It turns out that both estimators are strongly consistent and both converge to Gaussian processes. The covariance structures of these Gaussian processes are derived. An inconsistent MLE which is a self-consistent estimator is given in this section. All the proofs are in Section 4.

Before we close this introduction one should note that a different and more complicated competing-risk model is discussed by Dinse [3]. In view of our findings the nonparametric MLE, which is found through an EM algorithm [2] in Dinse's paper, may not be a consistent estimate.

2. Nonparametric Maximum Likelihood and Self-Consistency Approaches

To derive the nonparametric estimator we begin by assuming that the possible distributions F (the true failure time distribution) and G (the censoring distribution) are both discrete, with atoms $\{f_i\}$, $\{g_i\}$, respectively, at countably many specified points $0 = a_0 < a_1 < a_2 \ldots$, which usually include all observed values $\{Z_i = z_i\}_{i=1}^n$. Here we use the convention that all the distribution functions discussed in this paper are left continuous; i.e., if $X \sim F$ then $F(x) = P\{X < x\}$ and

$$\bar{F}(x) = P\{X \geqslant x\}.$$

Since any discrete survival function $\bar{F}(x)$ can be expressed in terms of the discrete hazard function h_j as

$$\bar{F}(x) = \prod_{x_j \leqslant x} (1 - h_j),$$

where $\{x_j\}$ denote the support of F,

$$h_j = \lim_{\varepsilon \to 0^+} \frac{P(X \in [x_j, x_j + \varepsilon) \mid x_j \leqslant X)}{\varepsilon}.$$

The nonparametric (NP) MLE of $\bar{F}(x)$ is obtained simply by replacing h_j by $\hat{h}_j$, the MLE of h_j for all j.

To find the NP MLE of h_j, we first write down the full likelihood under the current situation,

$$LIK = \prod_{i=1}^{n} \{p[f(z_i)\, \bar{G}(z_i)]^{\delta_i}\, [\bar{F}(z_i^+)\, g(z_i)]^{1-\delta_i}\}^{\Delta_i}$$

$$\times \{(1-p)[f(z_i)\, \bar{G}(z_i) + \bar{F}(z_i^+)\, g(z_i)]\}^{1-\Delta_i},$$

where

$$\bar{F}(z_i^+) = 1 - F(z_i^+) = P\{X > z_i\} = 1 - \sum_{a_j \leqslant z_i} f(a_j).$$

In terms of the hazard function the likelihood function can be written

$$\begin{aligned} LIK = \prod_{i=1}^{n} & \left\{ p \left[h(z_i) \prod_{a_j < z_i} (1 - h(a_j))(1 - h'(a_j)) \right]^{\delta_i} \right. \\ & \left. \times \left[(1 - h(z_i))\, h'(z_i) \prod_{a_j < z_i} (1 - h(a_j))(1 - h'(a_j)) \right]^{1 - \delta_i} \right\}^{\Delta_i} \\ & \cdot \left\{ (1 - p)[h(z_i) + (1 - h(z_i))\, h'(z_i)] \prod_{a_j < z_i} (1 - h(a_j))(1 - h'(a_j)) \right\}^{1 - \Delta_i} \\ = p^{\sum_{i=1}^{n} \Delta_i} (1 - p)^{\sum_{i=1}^{n} (1 - \Delta_i)} \cdot \prod_{i=1}^{n} & \left\{ \left[h(z_i) \prod_{a_j < z_i} (1 - h(a_j))(1 - h'(a_j)) \right]^{\delta_i} \right. \\ & \left. \times \left[(1 - h(z_i))\, h'(z_i) \prod_{a_j < z_i} (1 - h(a_j))(1 - h'(a_j)) \right]^{1 - \delta_i} \right\}^{\Delta_i} \\ & \times \left\{ [h(z_i) + (1 - h(z_i))\, h'(z_i)] \prod_{a_j < z_i} (1 - h(a_j))(1 - h'(a_j)) \right\}^{1 - \Delta_i}. \end{aligned}$$

Here we use h_j' to denote the hazard function with respect to $\bar{G}$.

If there are $d(z_i)$ "observed failures," $c(z_i)$ "observed censorings," and $e(z_i)$ "uncertain cases" among $r(z_i)$ individuals in view at z_i, the contribution to the likelihood is (excluding p and $(1-p)$'s terms)

$$\begin{aligned} & h(z_i)^{d_i} (1 - h(z_i))^{c_i} h'(z_i)^{c_i} [h(z_i) + h'(z_i) - h(z_i)\, h'(z_i)]^{e_i} \\ & \quad \times [(1 - h(z_i))(1 - h'(z_i))]^{r_i - f_i}, \end{aligned}$$

where $d_i = d(z_i)$, $c_i = c(z_i)$, $e_i = e(z_i)$, and $f_i = d_i + c_i + e_i$.

By taking the derivatives of the log likelihood with respect to p, $h(z_i)$ and $h'(z_i)$ for all i and setting them to zero, we end up with the equations

$$\begin{aligned} & \frac{d_i}{h(z_i)} + \frac{e_i(1 - h'(z_i))}{h(z_i) + h'(z_i) - h(z_i)\, h'(z_i)} - \frac{r_i - (d_i + e_i)}{1 - h(z_i)} = 0 \\ & \frac{c_i}{h'(z_i)} + \frac{e_i(1 - h(z_i))}{h(z_i) + h'(z_i) - h(z_i)\, h'(z_i)} - \frac{r_i - f_i}{1 - h'(z_i)} = 0 \qquad (2.1) \\ & p = \frac{1}{n} \sum_{i=1}^{n} \Delta_i. \end{aligned}$$

There is no question about the MLE of P, which is $\hat{p}=(1/n)\sum_{i-1}^{n}\Delta_i$. However, the first two equations cannot uniquely determine $h(z_i)$ and $h'(z_i)$, in general. As a result, the NP MLEs are not unique.

The following theorem characterizes the solution set of all nonparametric MLEs when all $\{z_i\}$ are distinct. Let $z_{(i)}$ denote the ith ordered value of $\{z_i, z_2, ..., z_n\}$. Define a collection of distributions in pairs by

$$R=\left\{\begin{array}{l}(\hat{F},\hat{G});\ \hat{F}(x)=\prod_i{}^{*}\left(1-\frac{1}{r_i}\right)^{\delta_i}\prod_j{}^{**}\left(1-\frac{1}{r_j}\right)^{\psi_j}, \\ \qquad \hat{G}(x)=\prod_i{}^{*}\left(1-\frac{1}{r_i}\right)^{1-\delta_i}\prod_j{}^{**}\left(1-\frac{1}{r_j}\right)^{1-\psi_j}, 0\leqslant\psi_j\leqslant 1\end{array}\right\}, \tag{2.2}$$

The product "$\prod^*$" runs over all i such that $z_{(i)}<x$ if the corresponding $\delta_{(i)}$ is observed, the product "$\prod^{**}$" runs over all $z_{(j)}<x$ such that $\delta_{(j)}$ is missing ($\Delta_{(j)}=0$), and r_i denotes the size of the risk set in view of $z^{(i)}$. If $z^{(1)}<z_{(2)}<\cdots<z_{(n)}$ then $r_i=n-i+1$.

THEOREM 1. *Suppose that all* $\{z_i\}_{i=1}^n$ *are distinct (i.e.,* $f_i=1$ *for all* $1\leqslant i\leqslant n$*); then the solution set of all nonparametric MLEs of* F *and* G *coincides with the collection* R.

The proof of this theorem is deferred to Section 4. Since the NP MLEs are not unique, a natural question arises: "How do these estimators behave in general?" As we shall see later, some of them are consistent, but some of them are not.

The term "self-consistency" was introduced by Efron [6]. An estimator which is self-consistent is regarded as a promising estimator. We shall use this concept to find all self-consistent estimators. It turns out (Theorem 2) that the collection of all self-consistent estimators coincides with R again if all (z_j) are distinct.

To reflect the current situation, we will call a pair estimator $(\bar{\bar{F}}, \bar{\bar{G}})$ of $(\bar{F}, \bar{G})$ self-consistent if $\bar{\bar{F}}(z)\,\bar{\bar{G}}(z)=\#$ of $(z_i\geqslant z)/n$ for all z and the following two equations are satisfied:

$$\bar{\bar{F}}(z)=\frac{1}{n}\sum_{i=1}^{n} I(z_i\geqslant z)+\frac{1}{n}\sum_{\substack{z_i<z\\ \delta_i=0\\ \Delta_i=1}}\frac{\bar{\bar{F}}(z)}{\bar{\bar{F}}(z_i^+)}+\frac{1}{n}\sum_{\substack{z_i<z\\ \Delta_i=0}}\frac{\bar{\bar{F}}(z)\,\tilde{g}(z_i)}{\tilde{f}(z_i)\,\bar{\bar{G}}(z_i)+\bar{\bar{F}}(z_i^+)\,\tilde{g}(z_i)} \tag{2.3}$$

$$\bar{\bar{G}}(z)=\frac{1}{n}\sum_{i=1}^{n} I(z_i\geqslant z)+\frac{1}{n}\sum_{\substack{z_i<z\\ \delta_i=1\\ \Delta_i=1}}\frac{\bar{\bar{G}}(z)}{\bar{\bar{G}}(z_i)}+\frac{1}{n}\sum_{\substack{z_i<z\\ \Delta_i=0}}\frac{\tilde{f}(z_i)\,\bar{\bar{G}}(z)}{\tilde{f}(z_i)\,\bar{\bar{G}}(z_i)+\bar{\bar{F}}(z_i^+)\,\tilde{g}(z_i)}, \tag{2.4}$$

where $\tilde{g}(z_i)=\tilde{\bar{G}}(z_i)-\tilde{\bar{G}}(z_i^+)$ and $\tilde{f}(z_i)=\tilde{\bar{F}}(z_i)-\tilde{\bar{F}}(z_i^+)$. It should be noted that every term in the last summation of (2.3) gives an estimate for the conditional probability $P(X_i \geqslant z \mid Z_i = z_i, \Delta_i = 0)$.

THEOREM 2. *Suppose that all $(z_i)_{i=1}^n$ are distinct. Then the collection of all self-consistent estimators coincides with R.*

From Theorem 1 and Theorem 2, the collection R is the set of all NP MLEs as well as all self-consistent estimators. However, not all estimators in R are consistent. This fact will be shown at the end of next section. As a result, neither approach will lead to a consistent estimator in general. In view of this, we shall consider two estimators in the next section.

3. Two Estimators

As we claimed at the end of previous section the NP MLEs and "self-consistent" estimators are not consistent in general. We shall consider two alternative estimators in this section. It is shown that both estimators are strongly consistent. Furthermore, the asymptotic properties for both estimators are analyzed. As a result, the fact about the consistency of second estimator (see 3.3 below) provides a direct proof of the inconsistency of another estimator which is in R and hence a self-consistent and NP MLE. Define

$$\hat{\bar{F}}_A(z) = \prod_{\substack{z_j < z \\ \delta_j = 1 \\ \Delta_j = 1}} \left(1 - \frac{d_j}{r_j}\right) \cdot \prod_{\substack{z_j < z \\ \Delta_j = 0}} \left(1 - \frac{e_j}{r_j}\right)^{\psi(z_j)} \tag{3.1}$$

where d_j, e_j, r_j are defined as in Section 2, and

$$\psi(x) = \frac{\#\text{ of } \{z_j;\, z_j \geqslant x,\, \delta_j = 1,\, \Delta_j = 1\}}{\#\text{ of } \{z_j;\, z_j \geqslant x,\, \Delta_j = 1\}},$$

a natural estimate of the ratio of being a failure based on the knowledge $Z_i = X_i \wedge Y_i = x$ and $\Delta_i = 0$. The following theorem 3 provides a convenient way of analyzing the properties of $\hat{\bar{F}}_A$.

Let us define $\bar{H}(z) = \bar{F}(z)\,\bar{G}(z)$, $\bar{H}_{1\cdot}(z) = P\{Z \geqslant z,\, \delta = 1\} = \bar{H}_1(z)$, $\bar{H}_{11}(z) = P\{Z \geqslant z,\, \delta = 1,\, \Delta = 1\}$, $\bar{H}_{\cdot 1}(z) = P\{Z \geqslant z,\, \Delta = 1\}$, $\bar{H}_{10}(z) = P\{Z \geqslant z,\, \delta = 1,\, \Delta = 0\}$, and $\bar{H}_{\cdot 0}(z) = P\{Z \geqslant z,\, \Delta = 0\}$. Since (Z, δ) and Δ are independent, it follows that $\bar{H}_{11}(z) = p\bar{H}_{1\cdot}(z)$, $\bar{H}_{\cdot 1}(z) = p\bar{H}(z)$, $\bar{H}_{\cdot 0}(z) = (1-p)\,\bar{H}(z)$. Let T_F, T_G, and T_H denote the upper limit of the support of F, G, and H, respectively. Let $\hat{H}$, $\hat{H}_{1\bullet}$, $\hat{H}_{11}$, $\hat{H}_{\bullet 1}$, $\hat{H}_{10}$, and $\hat{H}_{\bullet 0}$ be the sample counterparts. Define

$$\xi(Z, \delta, \Delta, z) = \left[\int_0^{Z \wedge z} \frac{d\bar{H}_1(s)}{\bar{H}^2(s)} + \frac{I(Z < z, \delta = 1, \Delta = 1)}{p\bar{H}(Z)} \right.$$

$$- \left(\frac{1-p}{p}\right) \frac{\bar{H}_1(z)}{\bar{H}^2(Z)} I(Z < z, \Delta = 1)$$

$$+ \left(\frac{1-p}{p}\right) \int_0^t \frac{I(Z \geqslant s, \Delta = 1)}{\bar{H}} d\left(\frac{\bar{H}_1}{\bar{H}}\right)$$

$$\left. + \frac{\bar{H}_1(Z)}{\bar{H}^2(Z)} I(Z < z, \Delta = 0) - \int_0^z \frac{I(Z \geqslant s, \Delta = 0)}{\bar{H}(s)} d\left(\frac{\bar{H}_1}{\bar{H}}\right) \right] \bar{F}(z).$$

THEOREM 3. *Assuming that F is continuous, one can write*

$$\hat{\bar{F}}_A(z) - \bar{F}(z) = \frac{1}{n} \sum_{i=1}^{n} \xi(Z_i, \delta_i, \Delta_i, z) + r_n(z), \tag{3.2}$$

where

$$\sup_{0 \leqslant z \leqslant T} |r_n(z)| = O(n^{-3/4} (\log n)^{3/4}) \qquad a.s.\ for \quad T < T_H.$$

COROLLARY 1. *As an estimator of F, $\hat{F}_A(z)$ is uniformly strongly consistent on $[0, T]$, $T < T_H$, i.e.,*

$$\lim_{n \to \infty} \sup_{0 \leqslant z \leqslant T} |\hat{F}_A(z) - F(z)| = 0 \qquad a.s.$$

Furthermore,

$$n^{1/2}(\hat{F}_A(z) - F(z))$$

converges weakly to a mean zero Gaussian process on $D[0, T]$ with covariance structure

$$\Gamma_{z_1, z_2} = \operatorname{cov}(\xi(Z_1, \delta_1, \Delta_1, z_1), \xi(Z_1, \delta_1, \Delta_1, z_2))$$

$$= \bar{F}(z_1)\, \bar{F}(z_2) \left\{ \frac{1 - p + p^2}{p} \int_0^{z_1 \wedge z_2} [\bar{H}(s)]^{-2}\, dH_1(s) \right.$$

$$\left. - \frac{2(1-p)}{p} \int_0^{z_1 \wedge z_2} \bar{H}_1 [\bar{H}]^{-3}\, dH_1 + \frac{1-p}{p} \int_0^{z_1 \wedge z_2} \bar{H}_1^2 \bar{H}^{-4}\, dH \right\}.$$

We now define the second estimator

$$\hat{\bar{F}}_B(z) = \prod_{\substack{z_j < z \\ \delta_j = 1 \\ \Delta_j = 1}} \left(1 - \frac{d_j}{r_j}\right)^{1/\hat{p}}, \qquad \text{where} \quad \hat{p} = \frac{1}{n} \sum_i \Delta_i. \tag{3.3}$$

The estimator $\hat{\bar{F}}_B(z)$ is obtained by assigning heavier weights to all "observed failures." The properties of $\hat{\bar{F}}_B(z)$ can be found analytically through the following theorem.

THEOREM 4. *Assuming that F is continuous, one can write*

$$\hat{\bar{F}}_B(z) - \bar{F}(z) = \frac{1}{n} \sum_i \eta(Z_i, \delta_i, \Delta_i, z) + r_n'(z), \tag{3.4}$$

where

$$\eta(Z, \delta, \Delta, z) = \frac{\bar{F}(z)}{p} \left[\int_0^z \frac{I(Z \geqslant s)}{\bar{H}^2(s)} \, d\bar{H}_{11} + \frac{1}{\bar{H}(Z)} I(Z < z, \delta = 1, \Delta = 1) \right]$$
$$- \frac{\bar{F}(z)}{p^2} \left[(I(\Delta_i = 1) - p) \int_0^z \frac{d\bar{H}_{11}}{\bar{H}} \right],$$

and $\sup_{0 \leqslant z \leqslant T} |r_n'(z)| = O(n^{3/4} (\log \log n)^{3/4})$ *a.s.*

COROLLARY 2. *All the conclusions of Corollary 1 about $\hat{F}_A(z)$ still hold for $\hat{F}_B(z)$ except the covariance structure, which is given by*

$$\Gamma'_{z_i, z_2} = \operatorname{Cov}(\eta(Z_1, \delta_1, \Delta_1, z_1), \eta(Z_1, \delta_1, \Delta_1, z_2))$$
$$= \left\{ \frac{1}{p} \int_0^{z_1 \wedge z_2} \bar{H}^{-2} \, dH_1 - \frac{(1-p)}{p} \int_0^{z_1} \frac{dH_1}{\bar{H}} \int_0^{z_2} \frac{dH_1}{\bar{H}} \right\} \bar{F}(z_1) \bar{F}(z_2).$$

Assume both F and G are continuous from now on. Let $\hat{\bar{F}}_S(z) = \hat{\bar{F}}_B(z)^{\hat{p}}$. Since $\hat{\bar{F}}_B(z)$ is strongly consistent (from Corollary 2) it follows that $\hat{\bar{F}}_S(z)$ is not a consistent estimator of $\bar{F}(z)$ unless $p = 1$. If we let

$$\hat{\bar{G}}_s(z) = \prod_{\substack{z_j < z \\ \delta_j = 0 \\ \Delta_j = 1}} \left(1 - \frac{1}{r_j}\right) \cdot \prod_{\substack{z_i < z \\ \Delta_i = 0}} \left(1 - \frac{1}{r_i}\right),$$

it is easy to see $(\hat{F}_s, \hat{G}_s) \in R$. This shows that $(\hat{F}_s, \hat{G}_s)$, the self-consistent and NP MLE, is not consistent.

Remark. 1. $\hat{F}_B$ is neither a NP MLE nor self-consistent since there is no $\hat{G}_B$ such that $(\hat{F}_B, \hat{G}_B) \in R$.

2. No attempt has been made about the optimality of the proposed estimates. It seems to this author, in order to estimate the survival function efficiently, one needs to estimate the hazard function consistently. The methods of adaptive procedures developed recently may be useful in constructing an efficient estimator.

3. Another naive estimator can be made by discarding those data for which $\Delta_i = 0$ and constructing the usual product-limit estimator on the remaining data. Let m_n be the size of the remaining data. It is easy to see under the current randomly missing mechanism that $m_n/n \to p$ a.s. by the strong law of large numbers. If we use $\hat{F}_c(z)$ to denote this estimator, it is easy to show (see [10], for example) that $n^{1/2}(\hat{F}_c(z) - F(z))$ converges in distribution to a mean zero Gaussian process on $D[0, T]$ with covariance

$$\Gamma'' z_1, z_2 = \bar{F}(z_1)\,\bar{F}(z_2)\,\frac{1}{p}\int_0^{z_1 \wedge z_2} \bar{H}^{-2}\,dH_1.$$

In view of this, it is easy to see that the estimator $\hat{F}_B$ is more efficient than the naive estimator in the sense that the asymptotic variance of $n^{1/2}(\hat{F}_B(z) - F(z))$ is smaller than that of $n^{1/2}(\hat{F}_c(z) - F(z))$, and the net gain is $((1-p)/p)[\int_0^z \bar{H}^{-1}\,dH_1]^2$ for all $0 \leqslant z \leqslant T$.

4. Proofs

Proof of Theorem 1. If $f_i = 1$ for all $1 \leqslant i \leqslant n$, the first two equations of (2.1) reduce to one of the following cases:

(i) if $d_i = 1$,

$$\frac{1}{h(z_i)} - \frac{r_i - 1}{1 - h(z_i)} = 0 \Leftrightarrow h(z_i) = \frac{1}{r_i}$$

(ii) if $c_i = 1$,

$$\frac{1}{h'(z_i)} - \frac{r_i - 1}{1 - h'(z_i)} = 0 \Leftrightarrow h'(z_i) = \frac{1}{r_i}$$

(iii) if $e_i = 1$,

$$\frac{1 - h_i'}{h_i + h_i' - h_i h_i'} = \frac{r_i - 1}{1 - h_i} \Leftrightarrow h_i + h_1' = h_i h_i' + \frac{1}{r_i}$$

$$\Leftrightarrow h_i(1 - h_i') + h_i' = \frac{1}{r_i} \Rightarrow h_i \leqslant \frac{1}{r_i} \text{ and } h_i' \leqslant \frac{1}{r_i}$$

$$\Rightarrow \exists \psi_i \text{ such that } 0 \leqslant \psi_i = \frac{\log(1 - h_i)}{\log(1 - 1/r_i)} \leqslant 1.$$

It follows that

$$1-h_i=\left(1-\frac{1}{r_i}\right)^{\psi_i}, \qquad 1-h_i'=\left(1-\frac{1}{r_i}\right)^{1-\psi_i},$$

since $h_i+h_i'=h_ih_i'+1/r_i$. Combining (i), (ii), and (iii) we have shown that all maximum likelihood estimators must have the desired form (2.2).

Conversely, suppose that

$$\hat{\bar{F}}(x)=\prod_i{}^{*}\left(1-\frac{1}{r_i}\right)^{\delta_i}\prod_j{}^{**}\left(1-\frac{1}{r_j}\right)^{\psi_j}$$

$$\hat{\bar{G}}(x)=\prod_i{}^{*}\left(1-\frac{1}{r_i}\right)^{1-\delta_i}\prod_j{}^{**}\left(1-\frac{1}{r_j}\right)^{1-\psi_j}$$

for certain $\{\psi_j;\ 0\leqslant\psi_j\leqslant 1\}$. It suffices to show $\hat{\bar{F}}$, $\hat{\bar{G}}$ are the MLEs of $\bar{F}$ and $\bar{G}$.

Let $h_i=1/r_i$ and $h_i'=0$ if $d_i=1$; $h_i'=1/r_i$ and $h_i=0$ if $c_i=1$; $h_i=1-(1-1/r_i)^{\psi_i}$ and $h_i'=1-(1-1/r_i)^{1-\psi_i}$ if $e_i=1$. It is easy to check that the first two equations of (2.1) are satisfied, and the contribution to the likelihood in view of z_i is maximized by $(\hat{\bar{F}}, \hat{\bar{G}})$ for all three cases $d_i=1$, $c_i=1$, and $e_i=1$ (in fact, the likelihood function takes its maximum value (constant value) at all ψ_i, $0\leqslant\psi_i\leqslant 1$). The proof of this theorem is thus complete.

Proof of Theorem 2. We first show that any pair $(\tilde{F}, \tilde{G})$ in R must satisfy (2.3) and (2.4), and hence is self-consistent. Let

$$\hat{\bar{F}}(x)=\prod_i{}^{*}\left(1-\frac{1}{r_i}\right)^{\delta_i}\prod_j{}^{**}\left(1-\frac{1}{r_j}\right)^{\psi_j}$$

$$\hat{\bar{G}}(x)=\prod_i{}^{*}\left(1-\frac{1}{r_i}\right)^{1-\delta_i}\prod_j{}^{**}\left(1-\frac{1}{r_j}\right)^{1-\psi_j},$$

where $0\leqslant\psi_j\leqslant 1$ for all j such that $\Delta_j=0$.

It is clear that $(\hat{F}, \hat{G})\in R$. Let $z_{(1)}<z_{(2)}<\cdots<z_{(n)}$ be the ordered values of $\{z_i\}_{i=1}^n$. The proof proceeds via induction as follows.

If $z\leqslant z_{(1)}$, it is clear $(\hat{\bar{F}}(z), \hat{\bar{G}}(z))$ satisfy (2.3) and (2.4). Assuming $(\hat{\bar{F}}(z), \hat{\bar{G}}(z))$ satisfy (2.3) and (2.4) for $z\leqslant z_{(m)}$, we want to show that $(\hat{\bar{F}}(z), \hat{\bar{G}}(z))$ satisfy (2.3) and (2.4) for $z\leqslant z_{(m+1)}$. It suffices to show $(\hat{\bar{F}}(z_{(m+1)}), \hat{\bar{G}}(z_{(m+1)}))$ satisfy (2.3) and (2.4), since both $\hat{\bar{F}}(z)$, $\hat{\bar{G}}(z)$ are constant over the interval $(z_{(m)}, z_{(m+1)}]$.

Let us rewrite (2.3) and (2.4) as

$$\tilde{\tilde{F}}(z)=\frac{\sum_{i=1}^{n} I(z_i \geqslant z)}{n-\left[\sum_{\substack{z_i<z\\ \delta_i=0\\ \Delta_i=1}} \frac{1}{\tilde{\tilde{F}}(z_i^+)}+\sum_{\substack{z_i<z\\ \Delta_i=0}} \frac{g(z_i)}{f(z_i)\,\tilde{\tilde{G}}(z_i)+\tilde{\tilde{F}}(z_i^+)\,g(z_i)}\right]} \tag{4.1}$$

and

$$\tilde{\tilde{G}}(z)=\frac{\sum_{i=1}^{n} I(z_i \geqslant z)}{n-\left[\sum_{\substack{z_i<z\\ \delta_i=1\\ \Delta_i=1}} \frac{1}{\tilde{\tilde{G}}(z_i)}+\sum_{\substack{z_i<z\\ \Delta_i=0}} \frac{f(z_i)}{f(z_i)\,\tilde{\tilde{G}}(z_i)+\tilde{\tilde{F}}(z_i^+)\,g(z_i)}\right]}. \tag{4.2}$$

If $\Delta_{(m)}=1$ and $\delta_{(m)}=1$, it is easy to express

$$\begin{aligned}\hat{\hat{F}}(z_{(m+1)})=\hat{\hat{F}}(z_{(m)}^+)&=\hat{\hat{F}}(z_{(m)})\cdot\left(1-\frac{1}{r_{(m)}}\right)\\ &=\frac{r_{(m)}}{n-t_{(m)}}\left(\frac{r_{(m)}-1}{r_{(m)}}\right)=\frac{r_{(m)}-1}{n-t_{(m)}},\end{aligned}$$

where

$$t_{(m)}=n-\frac{r_{(m)}}{\hat{\hat{F}}(z_{(m)})}=\sum_{\substack{z_i<z_{(m)}\\ \delta_i=0\\ \Delta_i=1}} \frac{1}{\hat{\hat{F}}(z_i^+)}+\sum_{\substack{z_i<z_{(m)}\\ \Delta_i=0}} \frac{\hat{g}(z_i)}{\hat{f}(z_i)\,\hat{\hat{G}}(z_i)+\hat{\hat{F}}(z_i^+)\,\hat{g}(z_i)}.$$

It follows that $\hat{\hat{F}}(z)$ satisfies (4.1) for $z \in (z_{(m)}, z_{(m+1)}]$.

On the other hand, one can write

$$\hat{\hat{G}}(z_{(m+1)})=\frac{r_{(m)}-1}{n-(t'_{(m)}-1/r_{(m)}(n-t'_{(m)}))}=\frac{r_{(m)}}{n-t'_{(m)}}=\hat{\hat{G}}(z_{(m)})$$

where

$$t'_{(m)}=\sum_{\substack{z_i<z_{(m)}\\ \delta_i=0\\ \Delta_i=1}} \frac{1}{\hat{\hat{G}}(z_i)}+\sum_{\substack{z_i<z_{(m)}\\ \Delta_i=0}} \frac{\hat{f}(z_i)}{\hat{f}(z_i)\,\hat{\hat{G}}(z_i)+\hat{\hat{F}}(z_i^+)\,\hat{g}(z_i)}.$$

This shows that $\hat{\hat{G}}(z)$ satisfies (4.2) for $z \in (z_{(m)}, z_{(m+1)}]$.

With similar arguments, one can easily deduce that $(\hat{\hat{F}}(z), \hat{\hat{G}}(z))$ satisfy (4.1) and (4.2) when $z \in (z_{(m)}, z_{(m+1)}]$ if $\Delta_{(m)} = 1$ and $\delta_{(m)} = 0$. If $\Delta_{(m)} = 0$, one can write $\hat{\hat{F}}(z_{(m+1)})$ as

$$\hat{\hat{F}}(z_{(m+1)}) = \hat{\hat{F}}(z_{(m)}) \cdot \left(1 - \frac{1}{r_{(m)}}\right)^{\psi_{(m)}}$$

$$= \frac{r_{(m)}}{n - t_{(m)}} \cdot \left(1 - \frac{1}{r_{(m)}}\right)^{\psi_{(m)}},$$

since $\hat{\hat{F}}(z_{(m)})$ satisfies (4.1) by assumption

$$= \frac{r_{(m)} - 1}{n - t_{(m)}} \cdot \frac{(r_{(m)} - 1)^{\psi_{(m)} - 1}}{r_{(m)}^{\psi_{(m)} - 1}}$$

$$= \frac{r_{(m)} - 1}{n - t_{(m)}} \cdot \left(1 - \frac{1}{r_{(m)}}\right)^{-1 + \psi_{(m)}}.$$

Since

$$\hat{g}(z_{(m)}) = \left[1 - \left(1 - \frac{1}{r_{(m)}}\right)^{1 - \psi_{(m)}}\right] \hat{\hat{G}}(z_{(m)}),$$

$$\hat{f}(z_{(m)}) = \left[1 - \left(1 - \frac{1}{r_{(m)}}\right)^{\psi_{(m)}}\right] \hat{\hat{F}}(z_{(m)}),$$

it follows that

$$\frac{\hat{g}(z_{(m)})}{\hat{f}(z_{(m)})\, \hat{\hat{G}}(z_{(m)}) + \hat{\hat{F}}(z_{(m)}^{+})\, \hat{g}(z_{(m)})} = \frac{[1 - (1 - 1/r_{(m)})^{1 - \psi_{(m)}}]}{\hat{\hat{F}}(z_{(m)})(1/r_{(m)})}.$$

Therefore,

$$\bar{F}(\hat{z}_{(m+1)}) = \frac{r_{(m)} - 1}{(n - t_{(m)}) \cdot (1 - (1/r_{(m)}))^{1 - \psi_{(m)}}}$$

$$= \frac{r_{(m)} - 1}{n - t_{(m)} - [1 - (1 - (1/r_{(m)}))^{1 - \psi_{(m)}}]\, r_{(m)} \hat{F}^{-1}(z_{(m)})}$$

(note that $\hat{\hat{F}}(z_{(m)}) = r_{(m)}/n - t_{(m)})$ which is the solution of (4.1).

With the same argument, $\hat{\hat{G}}(z_{(m+1)})$ satisfies (4.2) when$'_{(m)} = 0$. This establishes the first half of the theorem.

Conversely, let $z_{(m)}$ be the smallest m such that $\Delta_{(m)} = 0$. This is equivalent to saying $\Delta_{(1)} = \Delta_{(2)} = \cdots = \Delta_{(m-1)} = 1$ and $\Delta_{(m)} = 0$. From a similar argument as that in Efron [6], one can show that if $(\tilde{\tilde{F}}, \tilde{\tilde{G}})$ satisfies (4.1) and (4.2), then $(\tilde{\tilde{F}}, \tilde{\tilde{G}})$ must have the form:

$$\left.\begin{aligned}\tilde{\tilde{F}}(z) &= \prod_i{}^* \left(1-\frac{1}{r_i}\right)^{\delta_i},\\ \hat{\tilde{G}}(z) &= \prod_i{}^* \left(1-\frac{1}{r_i}\right)^{1-\delta_i},\end{aligned}\right\} \quad \text{for} \quad z \leqslant z_{(m)}.$$

It follows that $\tilde{\tilde{F}}(z_{(m)})\,\tilde{\tilde{G}}(z_{(m)}) = (n-m+1)/n$.

From (2.3) and (2.4) we obtain

$$\tilde{\tilde{F}}(z^+_{(m)}) = \tilde{\tilde{F}}(z_{(m+1)}) = \tilde{\tilde{F}}(z_{(m)}) - \frac{1}{n} + \frac{1}{n}\,(\alpha),$$

$$\tilde{\tilde{G}}(z^+_{(m)}) = \tilde{\tilde{G}}(z^{(m+1)}) = \tilde{\tilde{G}}(z_{(m)}) - \frac{1}{n} + \frac{1}{n}\,(1-\alpha),$$

where

$$0 \leqslant \alpha = \frac{\tilde{\tilde{F}}(z_{(m+1)})\,\tilde{g}(z_{(m)})}{\tilde{f}(z_{(m)})\,\tilde{\tilde{G}}(z_{(m)}) + \tilde{\tilde{F}}(z_{(m+1)})\,\tilde{g}(z_{(m)})} \leqslant 1.$$

If we let $\tilde{\tilde{F}}(z^+_{(m)}) = \tilde{\tilde{F}}(z_{(m+1)}) = \tilde{\tilde{F}}(z_{(m)}) \cdot (1-1/(n-m))^{\psi(m)}$, it is clear that $\tilde{\tilde{F}}(z_{(m+1)})$ can be written as

$$\tilde{\tilde{F}}(z_{(m+1)}) = \tilde{\tilde{F}}(z_{(m)})\left[1 - \frac{1-\alpha}{n\tilde{\tilde{F}}(z_{(m)})}\right].$$

Therefore,

$$\left[1 - \frac{1-\alpha}{n\tilde{\tilde{F}}(z_{(m)})}\right] = \left(1 - \frac{1}{n-m}\right)^{\psi(m)} \Leftrightarrow$$

$$\psi_{(m)} = \log\left(1 - \frac{1-\alpha}{n\tilde{\tilde{F}}(z_{(m)})}\right) \cdot \left(\log\left(1 - \frac{1}{n-m}\right)\right)^{-1},$$

since $n\tilde{\tilde{F}}(z_{(m)}) \geqslant n\tilde{\tilde{F}}(z_{(m)})\,\tilde{\tilde{G}}_{(m)}) = n-m+1 > n-m$,

We obtain $0 \leqslant \psi_{(m)} \leqslant 1$. This shows that $\exists \psi_{(m)}$ between 0 and 1 such that

$$\tilde{\tilde{F}}(z_{(m+1)}) = \prod_i{}^* \left(1-\frac{1}{r_i}\right)^{\delta_i} \cdot \left(1-\frac{1}{n-m}\right)^{\psi(m)}.$$

Since $\tilde{\tilde{F}}(z_{(m+1)}) \cdot \tilde{\tilde{G}}(z_{(m+1)}) = (n-m)/n$, it follows that

$$\tilde{\tilde{G}}(z_{(m+1)}) = \prod_i{}^* \left(1-\frac{1}{r_i}\right)^{1-\delta_i} \cdot \left(1-\frac{1}{n-m}\right)^{1-\psi(m)}.$$

This shows that $(\tilde{\tilde{F}}, \tilde{\tilde{G}})$ has the desired from if $z \leqslant z_{(m+1)}$.

The rest of the proof proceeds by induction in a manner almost identical to the previous one. The proof of the theorem is thus complete.

Lemma 1 suffices for estimates of various probabilities used in proving Theorems 3 and 4. The proof is based upon the moment generating function and Markov's inequality. For a detailed proof see Chao and Lo [1].

LEMMA 1. *If $X_1, X_2, ..., X_n$ are i.i.d. with mean zero, $|X_i| \leqslant c$ and $\operatorname{Var}(X_i) = \sigma^2$, then for any positive z and d satisfying $cz \leqslant d$ and $nz\sigma^2 \leqslant d^2$ one has*

$$P\left(\left|\sum_{i=1}^{n} X_i\right| \geqslant 3d\right) \leqslant 2e^{-z}.$$

The next lemma provides the key estimates in various situations required for Theorem 3 and Theorem 4.

LEMMA 2. *If F is continuous, then*

$$\sup_{0 \leqslant z \leqslant T} \left| \int_0^z (\hat{\bar{H}}^{-1} - \bar{H}^{-1})\, d(\hat{\bar{H}}^* - \bar{H}^*) \right| = O(n^{-3/4}(\log n)^{3/4}) \qquad a.s.$$

where $\hat{\bar{H}}^* - \bar{H}^*$ is chosen from any of the four functions

$$\{\hat{\bar{H}}_{11} - \bar{H}_{11},\ \hat{\bar{H}}_{\cdot 1} - \bar{H}_{\cdot 1},\ \hat{\bar{H}}_{1\cdot} - \bar{H}_{1\cdot},\ \hat{\bar{H}}_{\cdot 0} - \bar{H}_{\cdot 0}\}.$$

The proof of this lemma is exactly the same as given in Lo and Singh [10, Lemma 2].

LEMMA 3. *If F is continuous, then*

$$\sup_{0 \leqslant z \leqslant T} \left| \int_0^z (\hat{\bar{H}}^{-1} - \bar{H}^{-1})\, d(\bar{H}_{\cdot 0}(\hat{\bar{H}}_{11} \hat{\bar{H}}_{\cdot 1}^{-1} - \bar{H}_{11} \bar{H}_{\cdot 1}^{-1})) \right|$$
$$= O(n^{-3/4}(\log n)^{3/4}) \qquad \text{a.s.}$$

$$\sup_{0 \leqslant z \leqslant T} \left| \int_0^z (\hat{\bar{H}}^{-1} - \bar{H}^{-1})\, d(\bar{H}_{11} \bar{H}_{\cdot 1}^{-1}(\hat{\bar{H}}_{\cdot 0} - \bar{H}_{\cdot 0})) \right|$$
$$= O(n^{-3/4}(\log n)^{3/4}) \qquad a.s.$$

Proof.

$$\left| \int_0^z (\hat{\bar{H}}^{-1} - \bar{H}^{-1})\, d(\bar{H}_{\cdot 0}(\hat{\bar{H}}_{11} \hat{\bar{H}}_{\cdot 1}^{-1} - \bar{H}_{11} \bar{H}_{\cdot 1}^{-1})) \right|$$
$$\leqslant \left| \int_0^z (\hat{\bar{H}}^{-1} - \bar{H}^{-1})\, d(\bar{H}_{\cdot 0} \hat{\bar{H}}_{\cdot 1}^{-1}(\hat{\bar{H}}_{11} - \bar{H}_{11})) \right|$$
$$+ \left| \int_0^z (\hat{\bar{H}}^{-1} - \bar{H}^{-1})\, d(\bar{H}_{\cdot 0} \bar{H}_{11} \hat{\bar{H}}_{\cdot 1}^{-1} \bar{H}_{\cdot 1}^{-1}(\hat{\bar{H}}_{\cdot 1} - \bar{H}_{\cdot 1})) \right|$$
$$= A_z + B_z \qquad \text{(say)}.$$

Lemma 2 does not apply directly to A_z and B_z. However, the argument given in Lo and Singh [10, Lemma 2] can readily extend it with some additional effort to cover our case; i.e., the extra terms

$$\bar{H}_{\cdot 0}\hat{H}_{\cdot 1}^{-1} \qquad \text{and} \qquad \bar{H}_{\cdot 1}\bar{H}_{11}\hat{H}_{\cdot 1}^{-1}\bar{H}_{\cdot 1}^{-1}$$

appearing in A_z and B_z do not cause any trouble in the argument. We thus have

$$\sup_{0\leqslant z\leqslant T}|A_z| + \sup_{0\leqslant z\leqslant T}|B_z| = O(n^{-3/4}(\log n)^{3/4}) \qquad \text{a.s.}$$

With a similar argument applying to

$$\left|\int_0^z (\hat{H}^{-1}-\bar{H}^{-1})\, d(\bar{H}_{11}\bar{H}_{\cdot 1}(\hat{H}_{\cdot 0}-\bar{H}_{\cdot 0}))\right|,$$

we obtain ther desired result.

Proof of Theorem 1. Let $\tilde{\hat{H}}_1(z)=\hat{H}_{11}(z)+\hat{H}_{\cdot 0}(z)\cdot\hat{H}_{11}(z)/\hat{H}_{\cdot 1}(z)$. It is easy to see that $\log\hat{F}_A(z)=\int_0^z d\tilde{\hat{H}}_1(s)/\hat{H}(s)$. Let us write

$$\int_0^z\left(\frac{d\tilde{\hat{H}}_1}{\hat{H}}-\frac{d\bar{H}_1}{\bar{H}}\right)=\int_0^z\left(\frac{d\tilde{\hat{H}}_1}{\hat{H}}-\frac{d\hat{H}_{1\cdot}}{\hat{H}}\right)+\int_0^z\left(\frac{d\hat{H}_{1\cdot}}{\hat{H}}-\frac{d\bar{H}_{1\cdot}}{\bar{H}}\right)$$
$$=I(z)+II(z) \qquad \text{(say)}. \tag{4.3}$$

$I(z)$ can be further written as

$$I(z)=\int_0^z \hat{H}^{-1}d[\hat{H}_{\cdot 0}\hat{H}_{11}\hat{H}_{\cdot 1}^{-1}-\hat{H}_{10}]$$
$$=\int_0^z \hat{H}^{-1}d[\hat{H}_{\cdot 0}\hat{H}_{11}\hat{H}_{\cdot 1}^{-1}-\bar{H}_{\cdot 0}\bar{H}_{11}\bar{H}_{\cdot 1}^{-1}]-\int_0^z \hat{H}^{-1}d(\hat{H}_{10}-\bar{H}_{10})$$
$$=I_a(z)-I_b(z) \qquad \text{(say)}. \tag{4.4}$$

$I_a(z)$ can be further decomposed as

$$I_a(z)=\int_0^z \bar{H}^{-1}d[\hat{H}_{\cdot 0}(\hat{H}_{11}\hat{H}_{\cdot 1}^{-1}-\bar{H}_{11}\bar{H}_{\cdot 1}^{-1})]$$
$$+\int_0^z \bar{H}^{-1}d[\bar{H}_{11}(\hat{H}_{\cdot 0}-\bar{H}_{\cdot 0})\,\bar{H}_{\cdot 1}^{-1}]+R(z), \tag{4.5}$$

where $R(z)$ is the difference between $I_a(z)$ and the sum of two integrals in (4.5). In view of Lemma 3,

$$\sup_{0\leqslant z\leqslant T}|R(z)|=O(n^{-3/4}(\log n)^{3/4}) \qquad \text{a.s.,}$$

it follows that up to a remainder of order $n^{-3/4}(\log n)^{3/4}$ (a.s.),

$$\begin{aligned}
I_a(z) &= \int_0^z \bar{H}^{-1} d[\bar{H}_{\bullet 0}\bar{H}_{11}(\hat{\bar{H}}_{\bullet 1}^{-1} - \bar{H}_{\bullet 1}^{-1}) + (\hat{\bar{H}}_{11} - \bar{H}_{11})\,\bar{H}_{\bullet 1}^{-1}] \\
&\quad + \int_0^z \bar{H}^{-1} d[(\hat{\bar{H}}_{\bullet 0} - \bar{H}_{\bullet 0})\,\bar{H}_{11}\bar{H}_{\bullet 1}^{-1}] \\
&= -\int_0^z \bar{H}^{-1} d(\bar{H}_{\bullet 0}\bar{H}_{11}\hat{\bar{H}}_{\bullet 1}\bar{H}_{\bullet 1}^{-2}) + \int_0^z \bar{H}^{-1} d(\bar{H}_{\bullet 0}\hat{\bar{H}}_{11}\bar{H}_{\bullet 1}^{-1}) \\
&\quad + \int_0^z \bar{H}^{-1} d[(\hat{\bar{H}}_{\bullet 0} - \bar{H}_{\bullet 0})\,\bar{H}_{11}\bar{H}_{\bullet 1}^{-1}] \\
&= I_{a1}(z) + I_{a2}(z) + I_{a3}(z) \qquad \text{(say).}
\end{aligned} \tag{4.6}$$

(Here we use Lemma 2, together with the fact that

$$\sup_{0 \leqslant z \leqslant T} |\hat{\bar{H}}_{\bullet 1}(z) - \bar{H}_{\bullet 1}(z)| = O(n^{-1/2}(\log n)^{1/2}) \text{ a.s.)}$$

One can rewrite I_{a1}, I_{a2}, I_{a3} as

$$\begin{aligned}
I_{a1}(z) &= \frac{1}{n}\sum_i \left\{ \frac{\bar{H}_{\bullet 0}(Z_i)\,\bar{H}_{11}(Z_i)}{\bar{H}_{\bullet 1}^2(Z_i)\,\bar{H}(Z_i)} I(Z_i < z, \Delta_i = 1) \right. \\
&\quad \left. - \int_0^z \frac{I(Z_i \geqslant s, \Delta_i = 1)}{\bar{H}(s)}\, d\left(\frac{\bar{H}_{\bullet 0}\,\bar{H}_{11}}{\bar{H}_{\bullet 1}^2}\right)\right\} \\
&= \left(\frac{1-p}{p}\right)\frac{1}{n}\sum_i \left\{\left[\frac{\bar{H}_1(Z_i)}{\bar{H}^2(Z_i)} I(Z_i < z, \Delta_i = 1)\right.\right. \\
&\quad \left.\left. - \int_0^z \frac{I(Z_i \geqslant s, \Delta_i = 1)}{\bar{H}}\, d\left(\frac{\bar{H}_1}{\bar{H}}\right)\right\}\right.
\end{aligned} \tag{4.7}$$

$$I_{a2}(z) = \left(\frac{1-p}{p}\right)\frac{1}{n}\sum_i \frac{I(Z_i < z, \delta_i = 1, \Delta_i = 1)}{\bar{H}(Z_i)} \tag{4.8}$$

$$\begin{aligned}
I_{a3}(z) &= \frac{1}{n}\sum_i \left[-\frac{\bar{H}_1(Z_i)}{\bar{H}^2(Z_i)} I(Z_i < z, \Delta_i = 0)\right. \\
&\quad \left. + \int_0^z \frac{I(Z_i \geqslant s, \Delta_i = 0)}{\bar{H}(s)}\, d\left(\frac{\bar{H}_1}{\bar{H}}\right) - \int_0^z \frac{(1-p)\, d\bar{H}_1}{\bar{H}}\right].
\end{aligned} \tag{4.9}$$

By Lemma 2, $II(z)$ can be written as (up to order of $n^{-3/4}(\log n)^{3/4}$ a.s.)

$$II(z) = \frac{1}{n}\sum_i \left[-\int_0^z \frac{I(Z_i \geqslant s)}{\bar{H}^2(s)}\, d\bar{H}_1 - \frac{1}{\bar{H}(Z_i)} I(Z_i < s, \delta_i = 1)\right]. \tag{4.10}$$

The theorem follows from (4.3) to (4.10) and two terms of Taylor's expansion of $\log \hat{\bar{F}} - \log \bar{F}$.

Proof of Corollary 1. Let $\xi_i(z) = \xi(Z_i, \delta_i, \Delta_i, z)$, $\bar{\xi}(z) = (1/n) \sum_{i=1}^n \xi_i(z)$. Some tedious calculations show that

$$E\xi_i(z) = 0 \qquad \text{and} \qquad \xi_{z_i,\, z_2} = \operatorname{cov}(\xi_1(z_1), \xi_1(z_2))$$

equal the desired quantity stated in the corollary. As a result it follows that, for any $z_1, z_2 \in [0, T]$,

$$\operatorname{Var}(\xi_1(z_1) - \xi_1(z_2)) \leqslant |V(z_1) - V(z_2)| \tag{4.11}$$

for a continuous nondecreasing function $V(z) = \text{const} \cdot (F + H_1)$. (Note that the continuity of F entails the continuity of H_1 here.)

The weak convergence of the process $n^{1/2}(\hat{F}_A - F)$ follows from the tightness arguments (see Billingsley, for example): for any $0 \leqslant z_1 \leqslant z_2 \leqslant T$,

$$\begin{aligned} &E\{n(\bar{\xi}(z_1) - \bar{\xi}(z))^2\, n(\bar{\xi}(z) - \bar{\xi}(z_2))^2\} \\ &\quad \leqslant \text{const} \cdot E(\xi_1(z_1) - \xi_1(z))^2 \;\; E(\xi_1(z) - \xi_1(z_2))^2 \\ &\quad \leqslant \text{const} \cdot (V(z_2) - V(z_1))^2. \end{aligned} \tag{4.12}$$

To establish the uniformly strong consistency, one can even show a stronger result as follows:

By standard arguments (see Proposition 2.1 in [11]), one can show that

$$\sup_{0 \leqslant z \leqslant T} |\hat{F}_A(z) - F(z)| = O(n^{-1/2}(\log\log n)^{1/2}) \qquad \text{a.s.},$$

a consequence of standard functional LIL. In this approach to functional LIL, one first establishes the finite dimensional LIL for $n^{1/2}\bar{\xi}/(\log\log n)^{1/2}$ which is easy in view of Theorem 1. To complete the proof one needs a bound to control the fluctuations of the process $n^{1/2}\bar{\xi}/(\log\log n)^{1/2}$. Such a bound is available and provided by (4.11).

Proof of Theorem 4. Since F is continuous, d_j must be at most 1 with probability 1 for all j. It follows that

$$\begin{aligned} \left|\log \hat{\bar{F}}_B(z) - \frac{1}{\hat{p}} \int_0^z \frac{d\hat{\bar{H}}_{11}}{\hat{\bar{H}}}\right| &\leqslant \sum\nolimits^* \left|\frac{1}{\hat{p}} \log\left(\frac{n-i}{n-i+1}\right) - \frac{1}{\hat{p}} \cdot \frac{1}{n} \cdot \frac{1}{\hat{\bar{H}}(Z_{(i)})}\right| \\ &\leqslant \frac{1}{\hat{p}} \sum\nolimits^* \left|\log\left(1 - \frac{1}{n-i+1}\right) + \frac{1}{n} \cdot \frac{n}{n-i}\right| \\ &= O(n^{-1}) \qquad \text{a.s. uniformly on } 0 \leqslant z \leqslant T, \end{aligned}$$

where $\sum^*$ runs over all $Z_{(1)} \leqslant Z_2) \cdots < z$ such that $\delta_{(i)} = 1$ and $\Delta_{(i)} = 1$, and the integer i appearing in $\sum^*$ is less than $(1-\varepsilon)n$ for a positive ε and for all n a.s.

We then write

$$\begin{aligned} D_z &= \frac{1}{\hat{p}} \int_0^z \frac{d\hat{H}_{11}}{\hat{H}} - \int_0^z \frac{d\bar{H}_{1\bullet}}{\bar{H}} = \frac{1}{\hat{p}} \int_0^z \frac{d\hat{H}_{11}}{\hat{H}} - \frac{1}{p} \int_0^z \frac{d\bar{H}_{11}}{\bar{H}} \\ &= \left[\frac{1}{\hat{p}} \int_0^z \frac{d\hat{H}_{11}}{\hat{H}} - \frac{1}{\hat{p}} \int_0^z \frac{d\bar{H}_{11}}{\bar{H}}\right] + \left[\frac{1}{\hat{p}} - \frac{1}{p}\right] \int_0^z \frac{d\bar{H}_{11}}{\bar{H}} \\ &= \frac{1}{\hat{p}} \left\{ \int_0^z (\hat{H}^{-1} - \bar{H}^{-1})\, d\bar{H}_{11} + \int_0^z \bar{H}^{-1} d(\hat{H}_{11} - \bar{H}_{11}) \right. \\ &\quad \left. + \int_0^z (\hat{H}^{-1} - \bar{H}^{-1})\, d(\hat{H}_{11} - \bar{H}_{11}) \right\} - \left[\frac{\hat{p} - p}{\hat{p}p}\right] \int_0^z \frac{d\bar{H}_{11}}{\bar{H}}. \end{aligned}$$

In view of Lemma 2 and the fact that $\hat{p} - p = O(n^{-1/2}(\log n)^{1/2})$ a.s. and $\sup |\hat{H} - \bar{H}| = O(n^{-1/2}(\log n)^{1/2})$ a.s.,

$$\begin{aligned} D_z &= \frac{1}{np} \sum_{i=1}^n \left[\int_0^z \frac{I(Z_i \geqslant s)}{\bar{H}^2(s)}\, d\bar{H}_{11} + \frac{1}{\bar{H}(Z_i)} I(Z_i \leqslant z,\, \delta_i = 1,\, \Delta_i = 1) \right] \\ &\quad - \frac{1}{np^2} \sum_{i=1}^n [I(\Delta_i = 1) - p] \cdot \int_0^z \frac{d\bar{H}_{11}}{\bar{H}} + O(n^{-3/4}(\log n)^{3/4}) \end{aligned}$$

uniformly on $[0, T]$ a.s.

The theorem follows from the expression of D_z and two terms of Taylor's expansion of $\log \hat{\bar{F}}_B - \log \bar{F}$.

Proof of Corollary 2. This corollary can be proved in a manner similar to that of Corollary 1, except that in the current case we obtain a different covariance structure

$$\Gamma'_{z_1, z_2} = \left\{ \frac{1}{p} \int_0^{z_1 \wedge z_2} \bar{H}^{-2}\, dH_1 - \frac{(1-p)}{p} \left[\int_0^{z_1} \frac{dH_1}{\bar{H}} \right] \left[\int_0^{z_2} \frac{dH_1}{\bar{H}} \right] \right\} \bar{F}(z_1)\, \bar{F}(z_2).$$

Acknowledgments

I should like to express my appreciation to Professor H. Chernoff, S. Lagakos, and Dr. C. H. Chen for helpful discussions and suggestions. The constructive comments made by a referee have improved the presentation of this work. The naive estimator discussed in Remark 3 was suggested by the referee.

REFERENCES

[1] CHAO, M. T., AND LO, S.-H. (1988). Some representations of the nonparametric maximum likelihood estimators with truncated data. *Ann. Statist.* **16** 661–668.

[2] DEMPSTER, A. P., LAIRD, N. M., AND RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Roy. Statist. Soc. Ser. B* **39** 1–38.

[3] DINSE, G. E. (1982). Nonparametric estimation for partially-complete time and type of failure data. *Biometrics* **38** 417–431.

[4] DINSE, G. E. (1985). Estimating disease prevalence and mortality when cause of death is uncertain for some animals. *Biometrics* **41** Abstract.

[5] DINSE, G. E. (1986). Nonparametric prevalence and mortality estimators for animal experiments with incomplete cause-of-death data. *J. Amer. Statist. Assoc.* **81** 328–336.

[6] EFRON, B. (1967). The two sample problem with censored data. In *Proceedings, Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4, (L. LeCam and J. Neyman, Eds), pp. 831–853. Univ. of California Press, Berkeley.

[7] KAPLAN, E. L., AND MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

[8] LAGAKOS, S. W. (1979). General right censoring and its impact on the analysis of survival data. *Biometrics* **35** 139–156.

[9] LAGAKOS, S. W. (1982). An evaluation of some two-sample tests used to analyze animal carcinogenicity experiments. *Utilitas Math.* **21** (Special Vol. B) 239–260.

[10] LO, S.-H., AND SINGH, K. (1986). The product limit estimator and the bootstrap: Some asymptotic representations. *Probab. Theory Related Fields* **71** 455–465.

[11] PHILIPS, W. (1977). A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables. *Ann. Probab.* **5** 319–350.

[12] RACINE-POON, A., AND HOEL, D. (1984). Nonparametric estimation of the survival function when cause of death is uncertain. *Biometrics* **40** 1151–1158.